Ritvik Shrivastava

ritvik03@gmail.com | +91 99 68 68 2004 Website | LinkedIn | Google Scholar

Education

Netaji Subhas Institute of Technology Delhi, India
Bachelor of Engineering – Information Technology June 2017
Honors: Department Rank 1 in junior and senior years, Academic Merit Scholar for 3 consecutive years
GPA: 8.8/10 *Junior/Senior GPA*: 9.3/10
Relevant Coursework: Neural Networks, Computer Vision, Advanced Calculus, Machine/Deep Learning (MOOC)

Professional Experience

Max Planck Institute for Software Systems Saarbrucken, Germany
Visiting Research Scholar (Fellowship) Dec 2017-Present

- Developing novel machine learning methods for fair classification of data in order to prevent discrimination of any kind against the actors involved.

IBM Research India Delhi, India
Research Intern May-Nov 2017

- Developed Hashtag generation and Stance Detection state-of-the-art systems using LSTMs and attention networks.
- Studied information flow, with focus on information diffusion, homophily and topic lifecycles in social networks.

Indian Institute of Technology, Delhi Delhi, India
Research Intern May 2016-April 2017

- Designed a novel paraphrase detection framework for both noisy text micro-blogging data and formal text.
- Constructed a stance detection model using sentiment & subjectivity analysis, resulting in unparalleled accuracies.

Publications

Conference Papers

1. **A Paraphrase and Semantic Similarity Detection System for User Generated Short-Text Content on Microblogs.**
   *The 26th International Conference on Computational Linguistics (**COLING**) 2016.* (PDF)
2. **Topical Stance Detection for Twitter: A Two-Phase LSTM Model Using Attention.**
   *The 40th European Conference on Information Retrieval (**ECIR**) 2018* (PDF)
3. **Topic Lifecycle on Social Networks: Analyzing the Effects of Semantic Continuity and Social Communities.**
   *The 40th European Conference on Information Retrieval (**ECIR**) 2018* (PDF)
4. **Twitter Stance Detection - A Subjectivity and Sentiment Polarity Inspired Two-Phase Approach**
   *International Conference on Data Mining (**ICDM**) 2017 SENTIRE Workshop.* (PDF)
5. **EmTaggeR: A Word Embedding Based Novel Method for Hashtag Recommendation on Twitter**
   *International Conference on Data Mining (**ICDM**) 2017 ACUMEN Workshop.* (PDF)
6. **A Big Data Analysis Framework Using Apache Spark and Deep Learning**
   *International Conference on Data Mining (**ICDM**) 2017 DSBDA Workshop.* (PDF)
7. **A Semantic Continuity Based Analysis of Topic Lifecycle on Social Networks**
   *The 6th International Conference on Complex Networks and their Applications (**COMPLEX NETWORKS**) 2017* (PDF)
8. **Assessing the Effects of Social Familiarity and Stance Similarity in Interaction Dynamics**
   *The 6th International Conference on Complex Networks and their Applications (**COMPLEX NETWORKS**) 2017* (PDF)
9. ***SemTagger: A Novel Approach for Semantic Similarity Based Hashtag Recommendation on Twitter***
   *The 14th International Conference on Natural Language Processing (**ICON**) 2017* (PDF)

Journal Paper

10. **Multi-Class Instance-Incremental Framework for Classification in Fully Dynamic Graphs**
    *International Journal of Computational Science and Engineering (**IJCSE**).* (PDF)

Book Chapter

11. **Knowledge Discovery: Temporal Disaggregation in Social Interaction Data**
    In book: *Spatio-Temporal Graph Data Analytics.* (Book Link) (Chapter PDF)

Skills

- **Programming Languages**: Python, Java, C/C++
- **Frameworks and Tools**: PyTorch, Weka, Stanford-CoreNLP, NLTK, Mallet, LibSVM, Noah's Ark – TweetNLP

Research Projects

Stance Detection using Attentional Bi-LSTM Neural Network Sep-Nov 2017
IBM Research India Delhi, India

- Devised a deep model for stance detection in Twitter tweets with respect to a given target phrase.
- Designed a two-phase Bi-LSTM based neural net, with a global attention channel for target phrase embedding.

Topic Lifecycle Analysis in Social Network Communities Aug-Oct 2017
IBM Research India Delhi, India

- Analyzed the variations in topic usage over time to estimate the lifecycle patterns of topics inside communities.
- Investigated the morphing of topics into semantically-similar variants over time within communities and as a result observed the influence of users on the communication patterns and semantics within a group.

Studying Information Flow and Homophily in Social Networks June-Oct 2017
IBM Research India Delhi, India

- Investigated the relationship between familiarity of users and textual similarity of their social media content, at the user, peer-group and community granularities, leading to a manuscript **under review at ICWSM 2018.**
- Studied the effect on politeness and agreement between two users with respect to their stance towards a plethora of topics.

Hashtag Prediction Models in Social Networks May-June 2017
IBM Research India Delhi, India

- Conceptualized three approaches to hashtag prediction in Twitter using machine learning, deep learning, topic modelling and word vectors, resulting in computationally light-weight frameworks.
  - Generated a two-stage hashtag prediction pipeline using LDA topic modelling (MALLET) and our benchmark paraphrase model to produce a lift of 6 times over the baseline systems in the first approach.
  - Modelled a second approach solely using word vectors, to develop a novel algorithm for the prediction model resulting in a lift of approximately 7.5 times over the baseline.
  - Developed a state-of-the-art multi-class, multi-label framework for hashtag recommendation using R-CNN and attentional deep neural network architectures. Manuscript **under review at IJCAI 2018.**

Bachelor's Thesis Project – Big Data Classifier using Deep Learning Jan-May 2017
Netaji Subhas Institute of Technology, University of Delhi Delhi, India

- Researched the benefits of using ML in conjugation with big data, to develop an original classification framework.
- Assembled a three-stage pipeline using cascade learning, deep learning and Apache Spark, resulting in an appreciable improvement over existing structures with respect to accuracy, time-efficiency and scalability.

Stance Detection using Subjectivity and Sentiment Analysis Aug 2016-April 2017
Indian Institute of Technology (IIT) Delhi Delhi, India

- Constructed a two-stage subjectivity and sentiment analysis based stance detection system using hand-crafted ML(SVM) features, outperforming the models presented at SemEval 2016 Task 6, and the benchmark by 5%.

Paraphrase and Semantic Similarity Detection May-July 2016
Indian Institute of Technology (IIT) Delhi Delhi, India

- Designed a machine learning based paraphrase and semantic similarity detection system for both noisy microblogging text and formal text (Microsoft Corpus), resulting in a state-of-the-art model.

Identification and Analysis of 'Sleeper Cells' in Social Networks Jan 2016-May 2017
Indraprastha Institute of Information Technology (IIIT-Delhi) Delhi, India

- Conceptualized the idea of temporal paths, resulting in a novel clustering algorithm that is successfully able to detect clusters at different time-granularities in Social Networks.
- Devised novel systems using the temporal clustering algorithm to detect temporally-transient communities and to identify influential nodes at varying time granularities.
- Authored a book chapter in a peer-reviewed Springer book on this research.

Multi-class Incremental Classifier for Fully-Dynamic Graphs Aug 2015-July 2016
Netaji Subhas Institute of Technology, University of Delhi Delhi, India

- Enhanced the technique of Incremental Learning to improve the accuracies in graph classification.
- Constructed the classifier using Weisfeiler-Lehmann graph kernels in conjugation with SVMs with the data input in a streamed manner to obtain better results than batch-based state-of-the-art systems.